Title: Variation in the KAP6-1 gene in Chinese Tan sheep and associations with variation in wool traits

Authors: Jinzhong Tao, Huitong Zhou, Hua Gong, Zuoqing Yang, Qing Ma, Long Cheng, Wei Ding, Yingkang Li, Jon G.H. Hickford

PII: S0921-4488(17)30206-7
DOI: http://dx.doi.org/doi:10.1016/j.smallrumres.2017.08.001
Reference: RUMIN 5528

To appear in: *Small Ruminant Research*

Received date: 13-3-2017
Revised date: 3-7-2017
Accepted date: 1-8-2017

Please cite this article as: Tao, Jinzhong, Zhou, Huitong, Gong, Hua, Yang, Zuoqing, Ma, Qing, Cheng, Long, Ding, Wei, Li, Yingkang, Hickford, Jon G.H., Variation in the KAP6-1 gene in Chinese Tan sheep and associations with variation in wool traits.Small Ruminant Research http://dx.doi.org/10.1016/j.smallrumres.2017.08.001

ort Communication

Variation in the KAP6-1 gene in Chinese Tan sheep and associations with variation in wool traits

Jinzhong Tao[a,b], Huitong Zhou[b,c], Hua Gong[b,c], Zuoqing Yang[a], Qing Ma[d], Long Cheng[e], Wei Ding[d], Yingkang Li[d,*], Jon G.H. Hickford[b,c,*]

[a] Agricultural College, Ningxia University, Yinchuan, 750021, China.

[b] International Wool Research Institute, Gansu Agricultural University, Lanzhou 730070, China.

[c] Gene-Marker Laboratory, Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln 7647, New Zealand.

[d] Animal Science Institute of Ningxia Agriculture and Forestry Academy, Yinchuan, 750002, China.

[e] Faculty of Veterinary and Agricultural Sciences, University of Melbourne Dookie Campus, Victoria 3647, Australia.

* Corresponding author.

Email addresses: nxnkycx@163.com (Yingkang Li);

Jonathan.Hickford@lincoln.ac.nz (Jon G.H. Hickford)

Highlights:

- Four variants of *KRTAP6-1* were found in Tan sheep, including two new variants.
- A 57-bp deletion associated with increased MFD was not found in Tan sheep.
- Variation in *KRTAP6-1* associated with wool traits in Tan sheep.

Abstract

The keratin-associated protein KAP6-1 gene (*KRTAP6-1*) was investigated for its effect on wool traits in Chinese Tan sheep. Two previously identified *KRTAP6-1* variants (*A* and *B*) and two newly identified *KRTAP6-1* variants (*D* and *E*) were detected with frequencies of 28.9%, 57.1%, 10.3% and 3.7% for *A*, *B*, *D* and *E*, respectively. The *KRTAP6-1* variant *C*, was not found in the Chinese Tan sheep. Variant *D* was found to be associated with increased straightened fibre-length at birth, while variant *E* was found to be associated with increased straightened fibre-length, the number of crimps and the degree of crimping at approximately 35 days *post-partum* (conventionally called "Er-mao"). Variant *E* was also associated with increased wool fibre growth rate from birth to Er-mao. These results suggest that variation in ovine *KRTAP6-1* may affect wool traits in early life in Chinese Tan sheep.

Keywords: Chinese Tan sheep; crimp; KAP6-1 gene (*KRTAP6-1*); variation; wool traits

1. Introduction

The keratin associated proteins (KAPs) are a structural component of wool fibres and form a matrix that cross-links with the intermediate filaments (Powell and Rogers, 1997). KAPs can be classified into three broad groups on the basis of their amino acid composition: the high glycine-tyrosine (HGT; 35 - 60 mol% glycine and tyrosine) KAPs, the high sulphur (HS; ≤ 30 mol% cysteine) KAPs and the ultra-high sulphur (UHS; > 30 mol% cysteine) KAPs (Gong et al., 2016). Of these, the HGT-KAPs are predominantly found in the orthocortex of wool fibres and they vary in abundance both between and within breeds (Rogers, 2006; Gillespie, 1990). For example, wool from Lincoln sheep contains less than 1% by weight HGT-KAPs, while Merino wool has between 4% and 12% of HGT-KAPs by weight (Gillespie, 1990). Merino felting lustre (FL) mutant wool contains a low amount of HGT-KAPs (Gillespie and Darskus, 1971), and the expression of the HGT-KAP genes has been revealed to be down-regulated in FL mutant follicles (Li et al., 2009). This variation in the content of HGT-KAPs in wool fibres raises the question of whether these proteins affect the properties of wool.

Three HGT-KAP families have been identified in sheep: KAP6, KAP7 and KAP8 (Gong et al., 2016). KAP6 is a diverse KAP family with five genes identified to date (Zhou et al., 2016). Recently variation in one KAP6 gene (*KRATP6-1*) has been reported to affect wool fibre diameter-associated traits (Zhou et al., 2015).

Tan sheep are indigenous to China and they are recognised for producing long, curled or crimped, white wool by the early age of approximately 35 days. This is known as the age of "Er-mao" by Chinese people. Four wool crimp patterns have been described for wool at 'Er-mao", and these can be translated into "random curly", "soft and big curly", "walnut flower curly" and "Chinese string curly" (Tao et al., 2017). Variation in these traits affects the value of wool and lamb pelts in China. The genetic basis of these characteristics is currently unknown, hence the

objective of this study was to characterize *KRTAP6-1* in Chinese Tan sheep and investigate whether variation if present in the gene, affects wool traits of economic value in China.

2. Materials and methods

2.1. Sheep investigated and wool traits recorded

Five hundred and twenty-nine Chinese Tan lambs from 23 sire-lines were investigated. The majority of the lambs were born as singles and only 14 of them were born as twins. These twins were removed from the association analyses and only the 515 single lambs were analysed.

As most rural Chinese wool and pelt producers do not have access to internationally standardised wool testing methods, they have devised and used their own methods to measure the wool traits that they are interested in and value. These measures include: the straightened fibre-length (on the lambs shoulder or scapular) at birth, the straightened scapular fibre-length at Er-mao (approximately 35 days of age), the crimped scapular fibre length at Er-mao, the number of crimps along the fibre at Er-mao, the crimp frequency (the number of crimps per cm of length of straightened fibre) at Er-mao, the ratio of the straightened length to the crimped length at Er-mao (called the degree of crimping), and the wool fibre growth rate (expressed in mm/day) from birth to Er-mao.

In this study, fibre length was measured using an ordinary metric ruler and the number of crimps was counted by skilled farm staff. At birth, the straightened lengths of wool fibres on the shoulder were measured and the birth dates were recorded. At Er-mao, the straightened lengths on the shoulder, the crimped lengths of wool fibres from the shoulder and the numbers of crimps for wool fibres from the shoulder were measured. The measurement dates were recorded. Based on the above measurements, the crimp frequency and the ratio of the straightened length to the crimped length of wool fibres at Er-mao, precise age in days at Er-mao and the fibre growth rate from birth to Er-mao were calculated.

Sheep blood samples were collected onto FTA cards (Whatman BioScience, Middlesex, UK) and DNA for analysis was purified from 1.2 mm punches from the cards, using the procedure described by Zhou *et al*. (2006).

2.2. PCR amplification of KRTAP6-1 and SSCP analysis

Ovine *KRTAP6-1* was amplified using the polymerase chain reaction (PCR) primers described in Zhou *et al.* (2015). These primers are 5'-TCTACCCGAGAACAACCTC-3' and 5'-AGGCAAGTCTTTAGTAGGAC -3' and they were synthesised by Integrated DNA Technologies (Coralville, IA, USA). The amplification was undertaken using the purified genomic DNA on one punch of the FTA paper, 0.25 μM of each primer, 150 μM of each dNTP (Bioline, London, UK), 2.5 mM of Mg^{2+}, 0.5 U of Taq DNA polymerase (Qiagen, Hilden, Germany) and 1× reaction buffer supplied, in a total reaction volume of 20 μL.

The thermal profile for amplification consisted of 2 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C, with a final extension time of 5 min at 72 °C. This was undertaken in S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA). Amplicons were visualized by electrophoresis in 1% agarose (Bioline) gels, using 1 × TBE buffer containing 200 ng/mL of ethidium bromide.

PCR amplicons were genotyped using single-strand conformational polymorphism (SSCP) analysis. A 0.7 μL aliquot of each amplicon was mixed with 7 μL of loading dye (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene-cyanol) and after denaturation at 95 °C for 5 min, the samples were cooled rapidly on wet ice and loaded on 16 cm × 18 cm, 10.5% acrylamide: bisacrylamide (37.5:1) (Bio-Rad) gels with 1.5% glycerol. Electrophoresis was performed using Protean II xi cells (Bio-Rad), at 280 V for 22 h at 20 °C in 0.5 × TBE buffer. The gels were silver-stained by the method of Byun *et al*. (2009). PCR

amplicons of the three previously described variants of ovine *KRTAP6-1* (Zhou et al., 2015) were included as standards in each gel.

2.3. Sequencing of allele variants

PCR amplicons representative of SSCP patterns different to known *KRTAP6-1* variants (Zhou et al., 2015) were selected for DNA sequencing using a previously described approach (Gong et al., 2011). Briefly, a single band corresponding to the variant was excised as a gel slice from the polyacrylamide gel, macerated, and then used as a template for re-amplification with the original primers. This second amplicon was then sequenced in both directions at the DNA Sequencing Facility at Lincoln University, New Zealand.

2.4. Statistical analyses

All statistical analyses were performed using SPSS (version 22; IBM). Firstly, Pearson correlation coefficients were calculated to test the strength of the relationship between the wool traits. General Linear Models (GLMs) were then used to assess the effect of the presence of a particular *KRTAP6-1* variant in a sheep's genotype on the various wool traits, including straightened wool fibre length and crimping number at birth; straightened wool fibre length, crimped wool fibre length, crimping number, crimp frequency at Er-mao, the degree of crimping at Er-mao; and wool growth rate from birth to Er-mao. Sire was found to have an effect on all wool traits and was therefore factored into the GLMs. Birth weight was not found to affect wool traits except for the straightened wool length at birth and was therefore not fitted into the models except for the models testing straightened wool length at birth. Straightened wool length at birth was found to have an effect on all other wool traits, and it was therefore factored into the models. Age (in days) of lambs was also factored into the models for the Er-mao wool traits, as these wool traits were measured at different ages for different lambs. Gender was not factored into the models as it was not found to affect any wool trait. Only main effects were tested.

3. Results

3.1. Polymorphism of KRTAP6-1 in Chinese Tan sheep

SSCP analysis of PCR amplicons identified four alleles of *KRTAP6-1* in the Chinese Tan sheep, including two previously identified variants (*A* and *B*) and two new variants (*D* and *E*) that have not been described previously (Figure 1). Variant *C* (Zhou et al., 2015) was not found in these Tan sheep. Variant *E* was defined by a SNP that has not previously been described and was located upstream of the coding region. While variant *D* did not contain any new SNP, it had a sequence difference that appears to have resulted from recombination between variants *A* and *B*.

In total nine *KRTAP6-1* genotypes were detected in the 515 sheep. These genotypes and their frequencies were *AA* (9.3%), *AB* (30.5%), *AD* (6.4%), *AE* (2.1%), *BB* (35.0%), *BD* (10.0%), *BE* (3.9%), *DD* (1.6%) and *DE* (1.2%). This gives frequencies of 28.9%, 57.1%, 10.3% and 3.7% for variants *A*, *B*, *D* and *E*, respectively.

3.2. Correlations between wool traits

Strong correlations were found between many of the wool traits (Table 1).

3.3. Associations between KRTAP6-1 variation and wool traits

Associations were found between the four variants of *KRTAP6-1* found in the Tan sheep and variation in some of the wool traits measured for those sheep. These associations are summarized in Table 2.

4. Discussion

The identification of two additional variants of this gene takes the reported number of *KRTAP6-1* variants from three to five, suggesting *KRTAP6-1* is a polymorphic gene in sheep and confirming its effect on wool characteristics warrants further investigation.

Associations were revealed between variation in *KRTAP6-1* and various wool traits of value in Chinese markets. At Er-mao, a longer straightened fibre-length, a higher crimp number and a higher degree of crimping were found to be associated with variant *E*. While care is needed in accepting this finding, given that *E* is a less common variant, the possibility exists that the associations detected here may be due to the wool traits being correlated. For example, as the straightened fibre-length at Er-mao was moderately correlated with degree of crimping at Er-mao, the association of variant *E* with both traits, may be because of the correlation. In contrast, the associations obtained here cannot be solely explained by the wool trait correlations. For example, the degree of crimping at Er-mao was strongly correlated with crimped fibre-length at Er-mao, but there was not association between variant *E* and crimped fibre-length at Er-mao. Similarly, despite crimp number at Er-mao being strongly correlated with crimp frequency at Er-mao, no association was found between the absence/presence variant *E* and crimp frequency. This suggests *KRTAP6-1* may have independent effects on multiple wool traits.

Variation in *KRTAP6-1* has been reported to affect wool traits in Merino-cross sheep (Zhou et al., 2015). However, it is difficult to compare the results obtained from the Chinese Tan sheep, with those results. Firstly, different wool traits were measured for these two groups of sheep. Secondly, the variants (*D* and *E*) that were shown to be associated with wool traits in the Chinese Tan sheep, were not found in the Merino-cross sheep, whereas the variant (*C*) that was found to affect wool traits in the Merino-cross sheep, was not found in the Chinese Tan sheep. The absence of *KRTAP6-1* variant *C* in the Tan sheep studied is however interesting. This variant of *KRTAP6-1* has a 57-bp deletion in the coding region and this would result in the loss

of two GCGY repeats and an additional eleven amino acids in the central region of the protein. This variant has been reported to be associated with higher mean fibre diameter (MFD) wool in Merino-cross sheep (Zhou et al., 2015).

Tan sheep wool contains a mixture of two different types of fibres: the heterotype (containing medulation and non-medulation in the same fibre) and non-medulated fibres. Each type accounts for approximately 50% of the wool by weight. While the heterotype fibers are of a higher MFD (26.6 microns), the non-medulated fibres are fine with a mean diameter of 17.4 microns (Yang, 2011). The fineness of the non-medulated wool fibres in Chinese Tan sheep, would be consistent with the absence of variant *C* in this breed.

The effect of *KRTAP6-1* on wool growth in Chinese Tan sheep appears to be similar to the effect recently reported for another HGT-KAP gene called *KRTAP8-2* in this breed (Tao et al., 2017). *KRTAP6-1* and *KRTAP8-2* are clustered on sheep chromosome 1, together with other *KRTAPs*, including the genes encoding other KAP6 and KAP8 proteins, as well as the KAP7, KAP11, KAP13 and KAP24 families (Gong et al., 2012; Zhou et al., 2012; Gong et al., 2014; Zhou et al., 2016). Given that all these *KRTAPs* are polymorphic and potentially expressed (Gong et al., 2016), it would be difficult to determine whether the effect detected here is from variation in *KRTAP6-1* itself, or due to the linkage with variation in other *KRTAPs* that are nearby. Nevertheless the associations found in this study suggest that *KRTAP6-1* could be affecting wool traits in Tan sheep.

Conflict of Interest Statement

The authors have declared that no conflict of interest exists.

Dr Huitong Zhou

(for the authors)

Acknowledgements

The financial support from the Lincoln University Gene-Marker Laboratory and the Ningxia Special Breeding Program are acknowledged. We thank the support of the AGMARDT Postdoctoral Fellowship to Hua Gong, and the technical support from Seung Byun and Qian Fang.

References

Byun, S.O., Fang, Q., Zhou, H., Hickford, J.G.H., 2009. An effective method for silver-staining DNA in large numbers of polyacrylamide gels. Anal. Biochem. 385, 174-175.

Gillespie, J.M., 1990. The proteins of hair and other hard α- keratins, in: Goldman, R.D., Steinert, P.M. (Eds.), Cellular and Molecular Biology of Intermediate Filaments. Plenum Press, New York, pp. 95-128.

Gillespie, J.M., Darskus, R.L., 1971. Relation between the tyrosine content of various wools and their content of a class of proteins rich in tyrosine and glycine. Aust. J Biol. Sci. 24, 1189-1198.

Gong, H., Zhou, H., Hickford, J.G.H., 2011. Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (KAP6) family in sheep. Mol. Biol. Rep. 38, 31-35.

Gong, H., Zhou, H., Forrest, R.H.J., Li, S., Wang, J., Dyer, J.M., Luo, Y., Hickford, J.G.H., 2016. Wool keratin-associated protein genes in sheep – a review. Genes 7, 24.

Gong H., Zhou, H., Dyer, JM., Hickford, J.G.H., 2014. The sheep KAP8-2gene, a new KAP8 family member that is absent in humans. Springerplus 3, 528.

Gong, H., Zhou, H., McKenzie, G.W., Yu, Z., Clerens, S., Dyer, J.M., Plowman, J.E., Wright, M.W., Arora, R., Bawden, C.S., 2012. An Updated Nomenclature for Keratin-Associated Proteins (KAPs). Int. J Biol. Sci. 8, 258-264.

Li, S.W., Ouyang, H.S., Rogers, G.E., Bawden, C.S., 2009. Characterization of the structural and molecular defects in fibres and follicles of the merino felting lustre mutant. Exp. Dermatol. 18, 134-142.

Powell, B.C., Rogers, G.E., 1997. The role of keratin proteins and their genes in the growth, structure and properties of hair. In: Jollès, P., Zahn, H., Höcker, H. (Eds.), Formation and Structure of Human Hair, Birkhäuser Verlag, pp. 59-148.

Rogers, G.E., 2006. Biology of the wool follicle: an excursion into a unique tissue interaction system waiting to be rediscovered. Exp. Dermatol. 15, 931-949.

Tao, J., Zhou, H., Gong, H., Yang, Z., Ma, Q., Ding, W., Li, Y., Hickford, J.G.H., 2017. Variation in the KAP8-2 gene affects wool crimp and growth in Chinese Tan sheep. Small Rumin. Res. 149, 77-80.

Yang L., 2011. Polymorpism of the ovine KAP13 gene and its association with wool traits in Chinese Tan sheep. Master thesis. Ningxia University.

Zhou, H, Gong, H., Wang, J., Dyer, J.M., Luo, Y., Hickford, J.G.H., 2016. Identification of four new gene members of the KAP6 gene family in sheep. Sci. Rep. 6:24074.

Zhou, H., Gong, H., Li, S., Luo, Y., Hickford, J., 2015. A 57-bp deletion in the ovine KAP6-1 gene affects wool fibre diameter. J. Anim. Breed. Genet. 132, 301-307.

Zhou, H., Gong, H., Yan, W., Luo, Y., Hickford, J.G.H., 2012. Identification and sequence analysis of the keratin-associated protein 24-1 (KAP24-1) gene homologue in sheep. Gene 511, 62-65.

Zhou, H., Hickford, J.G.H., Fang, Q., 2006. A two-step procedure for extracting genomic DNA from dried blood spots on filter paper for polymerase chain reaction amplification. Anal. Biochem. 354, 159-161.

Figure Caption

Figr-1

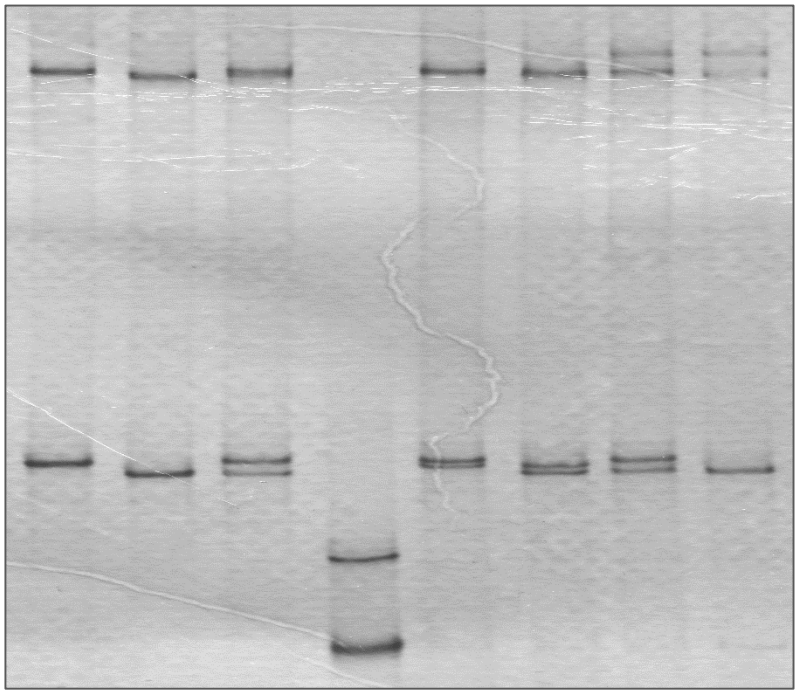

Figure 1. PCR-SSCP analysis of *KRTAP6-1* in selected sheep revealing five unique banding patterns. These were subsequently confirmed to be five unique DNA sequences (*A* to *E*) in the amplified region. The *CC* homozygous sheep was included for comparison, but this variant was not found in the Tan sheep in this study.

Table 1. Pearson correlation coefficients, *r*, were calculated to test the strength of the associations between the various wool traits[1]

	Straightened fibre-length at birth	Straightened fibre-length at Er-mao	Crimped fibre-length at Er-mao	Crimp number at birth	Crimp number at Er-mao	Crimp frequency at Er-mao
Straightened fibre-length at Er-mao	0.269 <0.001					
Crimped fibre-length at Er-mao	0.534 <0.001	0.094 0.032				
Crimp number at birth	**0.702** <0.001	0.234 <0.001	0.336 <0.001			
Crimp number at Er-mao	0.495 <0.001	0.259 <0.001	0.328 <0.001	**0.733** <0.001		
Crimp frequency Er-mao	0.284 <0.001	-0.397 <0.001	0.248 <0.001	0.543 <0.001	**0.778** <0.001	
Degree of crimping at Er-mao	-0.348 <0.001	0.468 <0.001	**-0.812** <0.001	-0.184 <0.001	-0.153 0.001	-0.444 <0.001

[1] For each pair of traits, the correlation coefficient (*r*) is shown in the top and the *P*-value is shown underneath it. Correlations with $|r| > 0.7$ are in bold, and those with $0.3 < |r| \leq 0.7$ are underlined.

Table 2. Associations between the presence/absence of *KRTAP6-1* variants and variation in wool traits (Mean ± SE)[1]

Trait	Variant assessed	n		Mean ± SE		*P*
		Absent	Present	Absent	Present	
Straightened fibre-length at birth (mm)	*A*	265	250	48.7 ± 0.7	48.1 ± 0.7	0.309
	B	107	408	48.5 ± 0.9	48.3 ± 0.7	0.833
	D	418	97	**48.1 ± 0.7**	**50.0 ± 0.9**	**0.010**
	E	477	38	48.4 ± 0.7	47.8 ± 1.2	0.575
Straightened fibre-length at Er-mao (mm)	*A*	265	250	88.2 ± 2.6	87.6 ± 2.6	0.620
	B	107	408	87.9 ± 2.8	88.0 ± 2.6	0.958
	D	418	97	88.4 ± 2.6	87.0 ± 2.8	0.376
	E	477	38	**87.2 ± 2.5**	**94.5 ± 3.3**	**0.002**
Crimped fibre-length at Er-mao (mm)	*A*	265	250	53.1 ± 0.7	53.3 ± 0.7	0.562
	B	107	408	53.3 ± 0.8	53.2 ± 0.7	0.826
	D	418	97	53.2 ± 0.7	53.2 ± 0.8	0.915
	E	477	38	53.3 ± 0.7	52.6 ± 0.9	0.330
Crimp number at birth (crimps)	*A*	265	250	5.8 ± 0.1	5.8 ± 0.1	0.512
	B	107	408	5.7 ± 0.2	5.8 ± 0.1	0.396
	D	418	97	5.8 ± 0.1	5.8 ± 0.1	0.682
	E	477	38	5.8 ± 0.1	5.7 ± 0.2	0.513
Crimp number at Er-mao (crimps)	*A*	265	250	6.3 ± 0.2	6.3 ± 0.2	0.658
	B	107	408	6.3 ± 0.2	6.3 ± 0.2	0.745
	D	418	97	6.3 ± 0.2	6.4 ± 0.2	0.525
	E	477	38	**6.3 ± 0.2**	**6.7 ± 0.3**	**0.020**
Crimp frequency Er-mao (crimps/cm)	*A*	265	250	0.82 ± 0.03	0.83 ± 0.03	0.453
	B	107	408	0.83 ± 0.03	0.83 ± 0.03	0.982
	D	418	97	0.82 ± 0.03	0.84 ± 0.03	0.232
	E	477	38	0.83 ± 0.03	0.81 ± 0.03	0.378
Degree of crimping at Er-mao	*A*	265	250	1.69 ± 0.06	1.68 ± 0.06	0.684
	B	107	408	1.69 ± 0.06	1.69 ± 0.05	0.994
	D	418	97	1.70 ± 0.06	1.67 ± 0.06	0.452
	E	477	38	**1.67 ± 0.05**	**1.82 ± 0.07**	**0.002**

[1]Predicted means, standard errors and *P* values from GLMs, with $P < 0.05$ being shown in bold..